**Study Information**

Hypotheses

Research question 1: machine learning models trained with work-related psychological constructs will show better predictive accuracy metrics when predicting turnover intention than models trained using biodata.

Research question 2: the predictive accuracy of machine learning models will increase when work-related psychological constructs are added to the model.

**Design Plan**

Study type

Observational Study - Data is collected from study subjects that are not randomly assigned to a treatment. This includes surveys, “natural experiments,” and regression discontinuity designs.

Blinding

* No blinding is involved in this study.

Is there any additional blinding in this study?

*No response*

Study design

This is not an experimental study. As there is no manipulation involved, it is neither a between-subject nor a within-subject design. See below for a description of the procedure.

*No files selected*

Randomization

The following publicly available instruments will be the instruments used for this study: the 3-items Utrecht Work Engagement Scale (UWES-3) (Schaufeli et al., 2017), the refined 9-items Lodahl and Kejner’s Job Involvement Scale (JI) (Reeve & Smith, 2001), the revised Three-Component Model (TCM) Employee Commitment Survey (Meyer & Allen, 1991), the 1-item Graphic Scale of Organizational Identification (GSOI) (Shamir & Kark, 2004), the 8-item Survey of Perceived Organizational Support (SPOS) (Shore & Tetrick, 1991), and the 6-items Big Five Iventory – Short Form (BFI-2-S). As our dependent variable we’ll use a one item Turnover Intention (TI) scale, which asks the following: “Are you contemplating leaving your organization within the next year?” Also as an dependented variable we’ll be using the Michigan Organizational Assessment Questionnaire (MOAQ). Additionally, biodata multiple choice questions will be asked, as well as demographic information.

All the participants will complete all the materials, and the order in which they take the scales will be the same for every participant.

**Sampling Plan**

Existing Data

Registration prior to creation of data

Explanation of existing data

*No response*

Data collection procedures

Participants for the study will be recruited from Amazon's Mechanical Turk (MTurk) to ensure a sample that closely reflects the diversity of the working population across various industries. The survey, hosted on MSU Qualtrics, will be distributed via Amazon's MTurk. MTurk participants will be paid $1.5 for participating in this study. Participants must be 18 years old and currently employed in order to qualify for participation in the study.

*No files selected*

Sample size

We aim to reach a target sample size of 250 participants for our study. To achieve this, we plan to recruit 250 participants from Amazon's Mechanical Turk (MTurk).

Sample size rationale

This study has been awarded an internal grant of $500 from the MSU I/O Psychology department. The allocated budget allows for compensating participants on MTurk at a rate of $1.5 per participant, limiting our data collection to 250 MTurk participants..

Stopping rule

*No response*

**Variables**

Manipulated variables

*No response*

*No files selected*

Measured variables

Engagement will be measured using the 3-items Utrecht Work Engagement Scale (UWES-3) (Schaufeli et al., 2017). Job involvement will be measured using the refined 9-items Lodahl and Kejner’s Job Involvement Scale (JI) (Reeve & Smith, 2001). Commitment will be measured using the affective and normative components of the revised Three-Component Model (TCM) Employee Commitment Survey (Meyer & Allen, 1991). Organizational identification will be measured using the 1-item Graphic Scale of Organizational Identification (GSOI) (Shamir & Kark, 2004). Perceived organizational support will be measured using the 8-item Survey of Perceived Organizational Support (SPOS) (Shore & Tetrick, 1991). Conscientiousness will be measured using the 6-items Big Five Inventory – Short Form (BFI-2-S). To measured turnover intention we’ll use a one item Turnover Intention (TI) scale, which asks the following: “Are you contemplating leaving your organization within the next year?” Finally, to measure job satisfaction we’ll be using the Michigan Organizational Assessment Questionnaire (MOAQ). Additionally, biodata multiple choice questions will be asked, as well as demographic information.

* [Qualtrics Survey - MTurk Participants.pdf](https://osf.io/project/ar78e/files/osfstorage/65bd3f75435c450602da74c3)
* [Qualtrics Survey - SONA Participants.pdf](https://osf.io/project/ar78e/files/osfstorage/65bd3f75435c450602da74c2)

Indices

We'll compute the mean of all items in the JSS scale, generating a consolidated measure of job satisfaction. Furthermore, we'll calculate the mean for each of the nine subscales in the JSS.

*No files selected*

**Analysis Plan**

Statistical models

Several machine learning algorithms will be used to train a series of models using biodata features only. These models includes GBT, RF, NN, SVM, and multiple regression. A second series of models that include work-related psychological constructs only in the data will be trained. Predictive accuracy metrics will be compared for each dataset to test hypothesis 3a. In addition, a baseline series of models will be trained using biodata features only. A second series of models that include work-related psychological constructs only in the data will be added to the models. Performance metrics of each model will be evaluated using accuracy, precision, recall, F1-scores, and AUC-ROC. Bootstrap resampling will be used as a statistical test to assess whether the improvement in performance when adding psychological constructs is statistically significant. An F-test will be used to determine if the change in R-squared is statistically significant for multiple regression. This is done to test for research question 2.

*No files selected*

Transformations

*No response*

Inference criteria

Our threshold for statistical significance will be set at an alpha value of p = 0.05.

Data exclusion

If any participants fail the attention check items, they will be excluded from the sample. Additionally, if the response time is lower than 1.5 minutes, the participants will also be excluded from the study. Outliers will be included in the analysis.

Missing data

*No response*

Exploratory analysis

Some exploratory analyses that we will conduct include employing different machine-learning algorithms to predict turnover intention and job satisfaction biodata and work-related psychological constructs as training data. Additionally, we will investigate any potential relationships of demographic variables with the MOAQ and TI scales. Demographic information will also be used to test for confounding effects in our results and potential sources of systematic bias in our data.

**Other**

Other

*No response*

[**Contributors**](https://osf.io/ar78e/contributors)

[Michael T. Bixter](https://osf.io/u8r2e) and [Diego](https://osf.io/ma3hs) Figueiras

**Description**

This study focuses on exploring whether work-related psychological constructs enhance ML models' predictive accuracy in turnover intention compared to biodata alone. The proposed hypotheses suggest that incorporating psychological constructs will improve predictive accuracy, addressing the "garbage in garbage out" concern prevalent in ML applications. The methods involve diverse datasets, including responses from federal employees and an online survey through Amazon's MTurk, with machine learning algorithms such as Gradient Boosting Trees, Random Forest, Neural Networks, Support Vector Machines, and logistic regression. The dissertation seeks to advance understanding in the field, offering practical insights for researchers and practitioners navigating the dynamic landscape of predictive modeling.

**Registration type**

OSF Preregistration

**Date registered**

February 2, 2024

**Date created**

February 2, 2024

**Associated project**

[osf.io/3tq2s](https://osf.io/3tq2s)

**Internet Archive link**

<https://archive.org/details/osf-registrations-ar78e-v1>

**Category**

 Project

**Registration DOI**

<https://doi.org/10.17605/OSF.IO/AR78E>

**Publication DOI**

No publication DOI

**Subjects**

* Psychology

* Social and Behavioral Sciences

* Quantitative Psychology

* Industrial and Organizational Psychology

**Affiliated institutions**

This registration has no affiliated institutions

**License**

CC0 1.0 Universal

**Tags**

No tags

**Citation**

osf.io/ar78e

Copyright © 2011-2024 [Center for Open Science](https://cos.io/)| [Terms of Use](https://github.com/CenterForOpenScience/centerforopenscience.org/blob/master/TERMS_OF_USE.md)| [Privacy Policy](https://github.com/CenterForOpenScience/centerforopenscience.org/blob/master/PRIVACY_POLICY.md)| [Status](https://status.cos.io/)| [API](https://developer.osf.io/)